

QY 7 LTGIVLVSVLLGVASHIDNYQPEQ 34

```

DB 465 LSTSTFGVSLVGLGVSSQIDNQPRLQ 492
RESULT 3
E88571
protein-G05B5-5 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
A:Accession: E88571
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99065613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/
A:Accession: E88571
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-585 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA83593.1; PID:g3873993; GSPDB:GNO0021; CESP:C05B5E
C:Genetics:
A:Gene: C05B5.5
A:Map position: 3

Query Match 29.0%; Score 60; DB 2; Length 585;
Best Local Similarity 46.4%; Pred. No. 5.3;
Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 7 LTGMVLVSVLLGVASHIDNYPPEQ 34
| | | | | | | | | | | | | | |
DB 465 LSTSTFGVSLVGLGVSSQIDNQPRLQ 492

RESULT 4
164215
hypothetical protein MG144 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999
A:Accession: 164215
R:Rieser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M. Fuhman, J.; Nguyen, D.; Uterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346; PMID:7569993
A:Accession: 164215
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-279 <TIGR>
A:Cross-references: GB:U39694; GB:U43667; NID:g1045822; PID:g1045827; TIGR:MG144
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3

Query Match 27.5%; Score 57; DB 2; Length 279;
Best Local Similarity 36.1%; Pred. No. 6.3;
Matches 13; Conservative 8; Mismatches 7; Indels 8; Gaps 1;

Qy 2 NRILTGMVLVSVLLGVA-----SHIDNY 29
: | | | | | | | | | | | | | | |
DB 231 SRSSTFGVSLVGLGVSSQIDNQPRLQ 266

RESULT 5
A55542
sensor kinase regeB - Rhodobacter capsulatus
C:Species: Rhodobacter capsulatus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
A:Accession: A55542
R:Mosley, C.S.; Suzuki, J.Y.; Bauer, C.E.
J. Bacteriol. 176, 7566-7573, 1994
A:title: Identification and molecular genetic characterization of a sensor kinase respon
sits.

```

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:35:59 ; Search time 11 Seconds
(without alignments)
154.594 Million cell updates/sec

Title: US-09-828-456-2

Perfect score: 207

Sequence: 1 MNRLIELTGWIVLVSVILL.....VASHIDNXPPEOSASVQHK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|----------------------|
| 1 | 207 | 100.0 | 41 | BLR_ECOLI | P56976 escherichia |
| 2 | 60 | 29.0 | 585 | YK05_CAEEL | P34293 caenorhabdit |
| 3 | 58.5 | 28.3 | 323 | C108_MOUSE | Q94773 mus musculus |
| 4 | 57 | 27.5 | 279 | Y144_MYCGE | P47390 mycoplasma |
| 5 | 55.5 | 26.8 | 55 | ATP8_PELSU | O79674 pelomonada |
| 6 | 55.5 | 26.8 | 426 | YAJ6_SCHPO | O09906 schizosacch |
| 7 | 52.5 | 25.4 | 120 | CHH1_PENUP | O15980 penaeus jap |
| 8 | 52.5 | 25.4 | 120 | CHH2_PENUP | O90562 penaeus jap |
| 9 | 52.5 | 25.4 | 572 | HENYA_PT3HV | P12364 human para |
| 10 | 52 | 25.1 | 124 | DHSC_R1CNC | O92199 rickettsia |
| 11 | 52 | 25.1 | 612 | OCTYC_BOVIN | O19094 bos taurus |
| 12 | 51.5 | 24.9 | 307 | O0X2_ACEAC | P50653 acetobacter |
| 13 | 51 | 24.6 | 458 | TCR_STRAG | P13924 streptococc |
| 14 | 50.5 | 24.4 | 104 | YHV2_LACHE | P22295 lactobacilli |
| 15 | 50.5 | 24.4 | 173 | CCME_NAEIN | P45036 haemophilus |
| 16 | 50.5 | 24.4 | 355 | RHOW_DROME | P20350 dirosophila |
| 17 | 50 | 24.2 | 219 | YDRE_SCHPO | O13748 schizosacch |
| 18 | 50 | 24.2 | 226 | ATP6_FELCA | P48994 felis silve |
| 19 | 50 | 24.2 | 265 | ASBP_HUMAN | P26436 homo sapien |
| 20 | 50 | 24.2 | 1127 | Y855_TREPA | O83827 treponema p |
| 21 | 49.5 | 23.9 | 404 | CYR_MARPO | P26852 marchantia |
| 22 | 49.5 | 23.9 | 513 | HUHV_STRPY | P58083 streptococc |
| 23 | 49 | 23.7 | 226 | ATP6_HIPAM | O94296 hippopotamu |
| 24 | 49 | 23.7 | 431 | PKR_RAV | O91787 tetanus toxin |
| 25 | 49 | 23.7 | 545 | ATPA_XENLA | P08428 xenopus lae |
| 26 | 49 | 23.7 | 1462 | TOR2_PEA | O24308 pisum sativ |
| 27 | 48.5 | 23.4 | 323 | C108_HUMAN | O94354 homo sapien |
| 28 | 48.5 | 23.4 | 360 | MRAY_PSEAE | O94928 pseudomonas |
| 29 | 48 | 23.2 | 66 | YP7B_TNVD | P27212 tobacco nec |
| 30 | 48 | 23.2 | 124 | DHSC_RICPR | P41085 rickettsia |
| 31 | 48 | 23.2 | 263 | RLPA_VIBCH | O9414 vibrio chol |
| 32 | 48 | 23.2 | 402 | KLRC_HUMAN | O9444 homo sapien |
| 33 | 48 | 23.2 | 503 | CUS_BACCHD | O94824 bacillus ha |

| | | | | | | |
|----|------|------|------|---|------------|--------------------|
| 34 | 47.5 | 22.9 | 392 | 1 | CYB_SOLTU | P29757 solanum tub |
| 35 | 47.5 | 22.9 | 393 | 1 | CYB_ARATH | P42792 arabidopsis |
| 36 | 47.5 | 22.9 | 394 | 1 | CYB_OENBE | P09843 oenothera b |
| 37 | 47.5 | 22.9 | 422 | 1 | EXUT_BACSU | O34456 bacillus su |
| 38 | 47.5 | 22.9 | 1953 | 1 | BIGA_SALTY | P25927 salmonella |
| 39 | 47 | 22.7 | 224 | 1 | UL01_HCVVA | P16719 human cytom |
| 40 | 47 | 22.7 | 242 | 1 | TRP2_CVACA | P34793 cyanidium c |
| 41 | 47 | 22.7 | 461 | 1 | FTSY_AQUAE | O67066 aquilex aeo |
| 42 | 47 | 22.7 | 462 | 1 | REGB_RHOSH | O53068 rhodobacter |
| 43 | 47 | 22.7 | 612 | 1 | OCTC_HUMAN | O94949 homo sapien |
| 44 | 47 | 22.7 | 1603 | 1 | VIT5_CAEEL | P06125 caenorhabdi |
| 45 | 46.5 | 22.5 | 203 | 1 | BCRC_BACLI | P42334 bacillus li |

ALIGNMENTS

RESULT 1
BLR_ECOLI
ID BLR_ECOLI STANDARD: PRT: 41 AA.
AC P56976;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-lactam resistance protein.
GN BLR OR B1624.1.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Butland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.D.,
RA Mau B., Shao Y., ~~Wong R.S.Y., McMurtry L.M., Levy S.B.;~~
RA ~~Wong R.S.Y., McMurtry L.M., Levy S.B.;~~
RT "intergenic" b1r gene in Escherichia coli encodes a 41-residue
RT membrane protein affecting intrinsic susceptibility to certain
RT inhibitors of peptidoglycan synthesis."
RL Mol. Microbiol. ~~7:354-362 (1990).~~
CC -1- FUNCTION: HAS AN EFFECT ON THE SUSCEPTIBILITY TO A NUMBER OF
CC ANTIBIOTICS INVOLVED IN PEPTIDOGLYCAN BIOSYNTHESIS. ACTS WITH BETA
CC LACTAMS, D-CYCLOSERINE AND BACITRACIN. HAS NO EFFECT ON THE
CC SUSCEPTIBILITY TO TETRACYCLINE, CHLORAMPHENICOL, GENTAMICIN,
CC FOSFOMYCIN, VACOMYCIN OR QUINOLONES. MIGHT ENHANCE DRUG EXIT BY
CC BEING PART OF MULTISUBUNIT EFFLUX PUMP. MIGHT ALSO BE INVOLVED IN
CC CELL WALL BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000258; -; NOT_ANNOTATED_CDS.
CC EMBL; AF219227; AAF82191.1; -.
CC Ecocore; EG14389; b1r.
KW Antibiotic resistance; Transmembrane; Inner membrane;
KW Complete proteome.
FT
FT TRANSMEM 4 24 POTENTIAL.

```

FT DOMAIN 25 41 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 41 AA; 4556 MM; D384EBB01276D46B CRC64;

Query Match
Best Local Similarity 100.0%; Score 207; DB 1; Length 41;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLEIETGWIIVLVSYILGVASHIDNYQPPQASVQHK 41
Db 1 MNRLEIETGWIIVLVSYILGVASHIDNYQPPQASVQHK 41

RESULT 2
YK05_CAEEL STANDARD; PRT; 585 AA.
AC p34293;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Hypothetical 66.5 kDa protein C0585.5 in chromosome III.
GN C0585.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Mortimore B.J.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z32679; CAA83593.1; -
DR WormRep; C0585.5; CE00051.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR002899; WRI/EB.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01683; EB; 2.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00289; WRI; 2.
DR PROSITE; PS50181; FBOX; 1.
DR Hypothetical protein.
KW DOMAIN
FT 12 60 F-BOX.
SQ SEQUENCE 585 AA; 66491 MM; E6951E9B854B8C69 CRC64;

Query Match
Best Local Similarity 29.0%; Score 60; DB 1; Length 585;
Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 7 LTGWIVLVSYILGVASHIDNYQPPQ 34
Db 465 LSTTFQIVLSLVGVSSQIDNDQPEIQ 492

RESULT 3
C108_MOUSE STANDARD; PRT; 323 AA.
AC 090Y73; 09D1P9; 0991Y8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein C10f8 homolog precursor (Thymic dendritic cell-derived factor
DE 1).
GN C10f8 OR TDCFL.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Thymus;
RA Jin C.G., Chen W.F.;
RT "Isolation and molecular cloning of gene encoding a novel dendritic
RT cell-derived factor.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stambli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G.,
RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Weiss C., Wittaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE OF 242-323 FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: LOCATIONS: Type I membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF116911; AAF20283.1; -
DR DR EMBL; AF003252; BAB22668.1; ALT_INIT.
DR EMBL; BC002164; AAH02164.1; -
DR MGD; MGI:1929278; Tdcfl.
KW Signal; Transmembrane.
FT 323 34
FT CHAIN 35 323
FT TRANSMEM 240 262
FT CARBOHYD 90 90
FT CONFLICT 48 48
FT CONFLICT 103 103
FT CONFLICT 121 121
FT CONFLICT 121 121
FT CONFLICT 139 139
FT CONFLICT 221 222
FT CONFLICT 264 264
FT CONFLICT 277 300
SQ SEQUENCE 323 AA; 36313 MM; 0E9BF6B6E07C7D96 CRC64;

Query Match
Best Local Similarity 28.3%; Score 58.5; DB 1; Length 323;
Matches 14; Conservative 9; Mismatches 4; Indels 9; Gaps 2;
```


RA Gabel C., Fuchs M., Filtz C., Holzer E., Moschl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambitt R., Purnelle B.,
 RA Goffeau A., Cadieu S., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Inoue M., Roemer M., Gattard C., Taitava V.R., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Domínguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of *Schistosoma mansoni* pombe".
 RL Nature 415:671-680(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE UPE0206 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z67961; CAA91892.1; -;
 DR InterPro: IPR005178; DUF300.
 DR Pfam: PF03619; DUF300; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 39 59 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 172 192 POTENTIAL.
 FT TRANSMEM 223 243 POTENTIAL.
 SQ SEQUENCE 426 AA; 49289 MW; 963450D799C3078 CRC64;
 Query Match 26.8%; Score 55.5; DB 1; Length 426;
 Best Local Similarity 28.9%; Pred. No. 10;
 Matches 11; Conservative 10; Mismatches 16; Indels 1; Gaps 1;
 QY 2 NRLELNGWLVVSVTLGVASHIDNYQPEQASVY 38
 DB 3 NEIVALCGFVIALVLCISITHLNKKPVLQRSV 40
 RESULT 7
 CHH2_PENUP STANDARD; PRT; 120 AA.
 AC 015980;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Crustacean hyperglycemic hormone 1 precursor (CHH) (Pej-SGP-I).
 OS Penaeus japonicus (Kuruma prawn).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata;
 OC Penaeoidea; Penaeidae; Marsupeneus.
 OX NCBI_Taxid=27405;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Eyestalk;
 RA Ohira T., Watanabe T., Nagasawa H., Aida K.;
 RT "Molecular cloning of cDNAs encoding four crustacean hyperglycemic
 RT hormones and a molt-inhibiting hormone from the kuruma prawn *Penaeus*
 RT *japonicus*".
 RL (in) Proceedings of the XIII International congress of comparative
 RL endocrinology, pp.83-86, Yokohama (1998).
 RP [2]
 RP SEQUENCE OF 47-118.
 RC TISSUE-Sinus gland;
 RX MEDLINE=97353924; Pubmed=9210164;
 RA Yang W.-J., Aida K., Nagasawa H.;
 RT "Amino acid sequences and activities of multiple hyperglycemic
 RT hormones from the kuruma prawn, *Penaeus japonicus*".
 RL Peptides 18:479-485(1997).

RN [3]
 RP SEQUENCE OF 47-101.
 RC TISSUE-Sinus gland;
 RA Yang W.-J., Aida K., Nagasawa H.;
 RT "Amino acid sequences of a hyperglycemic hormone and its related
 RT peptides from the kuruma prawn, *Penaeus japonicus*".
 RL Aquaticulture 135:205-212(1995).
 CC -1- FUNCTION: ABUNDANT HORMONE FOUND IN THE SINUS GLAND OF ISOPODS AND
 CC DECAPODS WHICH CONTROLS THE BLOOD SUGAR LEVEL. HAS A SECRETAGOGUE
 CC ACTION OVER THE AMYLASE RELEASED FROM THE MIDGUT GLAND. MAY ACT AS
 CC A STRESS HORMONE AND MAY BE INVOLVED IN THE CONTROL OF MOLTING AND
 CC REPRODUCTION.
 CC -1- TISSUE SPECIFICITY: PRODUCED BY THE MEDULLA TERMINALIS X-ORGAN IN
 CC THE EYESTALKS AND TRANSPORTED TO THE SINUS GLAND WHERE THEY ARE
 CC STORED AND RELEASED.
 CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD CHH/MIH/GIH/YIH FAMILY OF
 CC HORMONES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB007507; BAA22560.1; -;
 DR InterPro: IPR001166; CHH_MIH_GIH.
 DR Pfam: PF01147; Crust_neurohorm; 1.
 DR PRINTS: PRO0550; HYPERGLYCEMIC.
 DR PROSITE: PS01250; CHH_MIH_GIH; 1.
 KW Neuropeptide; Hormone; Glucose metabolism; Amidation; Signal;
 KW Cleavage on pair of basic residues; Multigene family.
 FT SIGNAL 1 24 POTENTIAL.
 FT PEPTIDE 25 44 CHH PRECURSOR RELATED PEPTIDE (CPRP).
 FT PEPTIDE 47 118 CRUSTACEAN HYPERGLYCEMIC HORMONE 1.
 FT DISULFID 53 89 BY SIMILARITY.
 FT DISULFID 69 85 BY SIMILARITY.
 FT DISULFID 72 98 BY SIMILARITY.
 FT MOD_RES 118 118 AMIDATION (G-119 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 120 AA; 13298 MW; CEC6BE04483897ED CRC64;
 Query Match 25.4%; Score 52.5; DB 1; Length 120;
 Best Local Similarity 41.4%; Pred. No. 7.1;
 Matches 12; Conservative 6; Mismatches 8; Indels 3; Gaps 1;
 QY 10 MIVLVSVILGV---ASHIDNYQPEQS 35
 DB 8 WSAIIASLILLLIAPASPYDAPSPPEAS 36
 RESULT 8
 CHH2_PENUP STANDARD; PRT; 120 AA.
 AC 0905D2; PB1681;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Crustacean hyperglycemic hormone 2 precursor (CHH) (Pej-SGP-II).
 OS Penaeus japonicus (Kuruma prawn).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata;
 OC Penaeoidea; Penaeidae; Marsupeneus.
 OX NCBI_Taxid=27405;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Eyestalk;
 RA Ohira T., Watanabe T., Aida K., Nagasawa H.;
 RT "Crustacean hyperglycemic hormone of kuruma prawn *Penaeus japonicus*".
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBD databases.
 RP [2]
 RP SEQUENCE OF 47-118.
 RC TISSUE-Sinus gland;

CC MEDLINE-97353924; PubMed-9210164;
CC Yang W.-J., Aida K., Nagasawa H.;
CC "Amino acid sequences and activities of multiple hyperglycemic
CC hormones from the kuruma prawn, Penaeus japonicus.";
CC Peptides 18:479-485(1997).
CC -1- FUNCTION: ABUNDANT HORMONE FOUND IN THE SWISS GLAND OF ISOPDS AND
CC DECAPDS WHICH CONTROLS THE BLOOD SUGAR LEVEL. HAS A SECRETAGOGUE
CC ACTION OVER THE AMYLASE RELEASED FROM THE MIDGUT GLAND. MAY ACT AS
CC A STRESS HORMONE AND MAY BE INVOLVED IN THE CONTROL OF MOLTING AND
CC REPRODUCTION.
CC -1- SIMILARITY: BELONGS TO THE APTRHOPOD CHH/MIH/GIH/VIH FAMILY OF
CC HORMONES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB035724; BAA88339.1; -
CC InterPro; IPR001166; CHH_MIH_GIH.
CC Pfam; PF01147; Crust_neurohorm. 1.
CC PRINTS; PR00550; HYPERGLYCEMIC. 1.
CC PROSITE; PS01250; CHH_MIH_GIH. 1.
CC Neuropeptide; Hormone; Glucose metabolism; Amidation; Signal;
CC Cleavage on pair of basic residues; Multigene family.
CC SIGNAL. 1
CC PEPTIDE 28 44 CHH PRECURSOR RELATED PEPTIDE (CPRP).
CC PERIPE 47 118 CRUSTACEAN HYPERGLYCEMIC HORMONE 2.
CC DISULFID 53 89 BY SIMILARITY.
CC FT DISULFID 69 85 BY SIMILARITY.
CC FT DISULFID 72 98 BY SIMILARITY.
CC MOD RES 118 118 AMIDATION (G-119 PROVIDE AMIDE GROUP).
CC SO SEQUENCE 120 AA; 13467 MW; C2D96499D1231585 CRC64;
CC
CC Query Match 25.4%; Score 52.5; DB 1; Length 120;
CC Best Local Similarity 41.4%; Pred. No. 7.1;
CC Matches 12; Conservative 6; Mismatches 8; Indels 3; Gaps 1;
CC
CC QY 10 WIVLVSVLLGV---ASHIDNKPPEQS 35
CC DB 8 WSAALLSLILLALLAPSVDAPSPPEAS 36
CC
CC RESULT 9
CC HEMA_P13HV
CC ID HEMA_P13HV STANDARD: PRT: 572 AA.
CC AC P12564;
CC DT 01-OCT-1989 (Rel. 12, Created)
CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hemagglutinin-neuraminidase (EC 3.2.1.18).
CC GN HN.
CC OS Human parainfluenza 3 virus (strain Tex/12677/83).
CC OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
CC OC Paramyxoviridae; Paramyxovirinae; Respirovirus.
CC OX NCBI_Taxid-11221;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE-88101361; PubMed-2827373;
CC RA van Wyke Coelingh K.L., Winter C.C., Murphy B.R.;
CC RT "Nucleotide and deduced amino acid sequence of hemagglutinin-
CC neuraminidase genes of human type 3 parainfluenza viruses isolated
CC from 1957 to 1983.";
CC VIROLOGY 162:137-143(1988).
CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
CC TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEURAMINIDASE
CC ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
CC THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
CC GLYCOPROTEINS.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,

CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -1- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
CC N-TERMINAL HYDROPHOBIC SEQUENCE.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES HEMAGGLUTININ-
CC NEURAMINIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M18764; AAA46851.1; -
CC PIR; P29970; HNNZ83.
CC InterPro; IPR000665; Hem-neuramidse.
CC DR Pfam; PF00423; HN; 1.
CC KM Hydrolase; Hemagglutinin; Envelope protein; Glycoprotein;
CC Transmembrane.
CC FT DOMAIN 1 31 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 32 53 POTENTIAL.
CC FT DOMAIN 54 572 EXTRACELLULAR (POTENTIAL).
CC FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 572 AA; 64330 MW; 57875F74AC568BED CRC64;
CC
CC Query Match 25.4%; Score 52.5; DB 1; Length 572;
CC Best Local Similarity 31.7%; Pred. No. 33;
CC Matches 13; Conservative 6; Mismatches 11; Indels 11; Gaps 1;
CC
CC QY 2 NRIELTGMIVLV-----VSIVLLGVASHIDNKP 31
CC DB 71 NEFMEITGRKIQMASDNTNDLIQSGVNRLLTIQSHVNYIP 111
CC
CC RESULT 10
CC DHSC_R1CCN
CC ID DHSC_R1CCN STANDARD: PRT: 124 AA.
CC AC Q92099;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Succinate dehydrogenase cytochrome b-556 subunit.
CC GN SDHC OR RC0168.
CC OS Rickettsia conorii.
CC OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
CC OC Rickettsiaceae; Rickettsiae; Rickettsia.
CC OX NCBI_Taxid=781;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Malish 7;
CC RX MEDLINE-21442074; PubMed-11557893;
CC RA Ogata H., Audic S., Renseto-Audiffren P., Fournier P.-E., Barbe V.,
CC RA Sanson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
CC RA Raoult D.;
CC RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
CC RL Science 293:2093-2098(2001).
CC -1- FUNCTION: MONO-HEME CYTOCHROME OF THE SUCCINATE DEHYDROGENASE
CC COMPLEX.
CC -1- PATHWAY: Fatty acid cycle.
CC -1- SUBUNIT: PART OF AN ENZYME COMPLEX CONTAINING FOUR SUBUNITS: A
CC FLAVOPROTEIN, AN IRON-SULFUR, CYTOCHROME B-556, AND AN HYDROPHOBIC
CC ANCHOR PROTEIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B560 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC used by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE008563; RALD02706-1;
DR InterPro; IPR000701; Sdh_cyt.
DR Pfam; PF01127; Sdh_cyt_1;
DR PROSITE; PS01000; Sdh_cyt_1; FALSE_NEG.
DR PROSITE; PS01001; Sdh_cyt_2; 1.
KW Tricarboxylic acid cycle; Electron transport; Home; Transmembrane;
KW Inner membrane; Complete proteome.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
FT BINDING 33 33 HOME (POTENTIAL).
FT BINDING 83 83 HOME (BY SIMILARITY).
SO SEQUENCE 124 AA; 58C6A0461AFCB1A C9C64;

| | | | | |
|-----------------------|--------|----------------|-------|---------------|
| Query Match | 25.1% | Score 52; | DB 1; | Length 124; |
| Best Local Similarity | 52.98% | Pred. NO. 8.5; | | |
| Matches | 9; | Conservative | 5; | Mismatches 3; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

```
QY      3 RLELTGWIVLWSVIL 19
      :|:| | | | | | |
Db     102 KLVNITGWCVVVGSVLL 118
```

| RESULT 11 | |
|---------------|-----------|
| OCTC_BOVIN | |
| ID OCTC_BOVIN | STANDARD; |
| | PRT; |
| | 612 AA |

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisomal carnitine octanoyltransferase (EC 2.3.1.-) (COT)

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea

| | |
|----|-------------------------------------|
| RN | [1] |
| RP | SEQUENCE FROM N.A., AND MUTAGENESIS |
| RC | TISSUE=Liver; |

RT "cDNA cloning, recombinant expression, and site-directed mutagenesis;
RT of bovine liver carnitine octanoyltransferase -- Arg505 binds the
RT carboxylate group of carnitine.";

CC CONCERN THE C6 TO C10 CHAIN LENGTH SUBSTRATE.
CC -1- PATHWAY: FATTY ACID BETA-OXIDATION CYCLE.
CC -1- SUBUNIT: MONOMER (PROBABLE).

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between EMBL, SIB, NIDDK and NCBI.
CC

CC SIMILARITY: BELONGS TO THE CARNITINE/CHOLINE ACETYLTRANSFERASE
CC FAMILY.

CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>)

CC -----
EMBL: U65745; AAC48758.1; -
DR InterPro: IPR000542; Carn_acyltrans
DR InterPro: IPR000966; Microb_diss C

```
DR   P1am; PF00755; Carn_acyltransf; 1.
DR   PROSITE; PS00439; ACYLTRANSF_C_1; 1
DR   PROSITE; PS00440; ACYLTRANSF_C_2; 1
```

| DR | PROSITE; PS00342; MICROBODIES_CTER; FALSE_NEG. |
|----|--|
| KW | Transferrase; Acyltransferase; Fatty acid metabolism; Transport. |
| KM | Peroxisome. |
| FT | ACT SITE 327 327 POTENTIAL. |
| ET | BINDING 505 505 INVOLVED IN SUBSTRATE (CARNITINE) |

| FT | SITE | 610 | 612 | MICROBODY TARGETING SIGNAL (POTENTIAL). |
|----|----------|---------|-----------|--|
| FT | MUTAGEN | 505 | 505 | R->N: INCREASE OF KM TOWARDS CARINITINE. |
| SQ | SEQUENCE | 612 AA: | 70263 MW; | 2D5D91A54CF8E2BA CRC64; |

| | | |
|-----------------------|--------|---------------|
| Best Local Similarity | 42.18; | Pred. NO. 40; |
| Matches | 8; | Mismatches |
| | 6; | Indels |
| | | Gaps |
| | | 0; |

```
QY 21 GVASHIDNYQPPEQASAVQ 39
    | |||::| ||:: ::
Db 105 GPASHIEHYWPPKEGTQLE 123
```

| | |
|---------------|-------------|
| RESULT 12 | |
| QOX2_ACEAC | |
| ID_QOX2_ACEAC | STANDARD; |
| | PRT; 307 AA |

| | |
|----|--|
| DT | 01-OCT-1996 (Rel. 34, last sequence update) |
| DT | 15-JUL-1999 (Rel. 38, last annotation update) |
| DE | ubiquinol oxidase polypeptide II precursor (EC 1.10.3.-) (Cytochrome |

05 Acetobacter aceti.
0C Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
0C Acetobacter.

RP SEQUENCE FROM N.A.
RC STRAIN-1023;
RX MEDLINE=93322308; Pubmed=8392509;

RT "Characterization of a cytochrome *al* that functions as a ubiquinol
oxidase in *Acetobacter aceti*.";
J. Bacteriol. 175:4307-4314(1993).

CC -1- SUBUNIT: HEHEHOLEKRAMER OF THE SUBUNITS 1', 2', 3' AND 4'.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: SOME, TO MITOCHONDRIAL OR BACTERIAL COX2 SUBUNITS
CC BUT LACK HEME-BINDING DOMAIN.

CC BY-NC-ND 4.0 International license.

```
CC CC entities requires a license agreement (see http://www.isd-sib.cn/announce)
CC or send an email to license@sb-sib.cn.
CC -----
DB ENMT D13195.PA02480.1. 2
```

```
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR ProDom; PD000071; Cytochrome_CuA_1
```

[illegible]

| | | | | |
|----|----------|---------|-----------|-------------------------|
| FT | TRANSMEM | 46 | 66 | POTENTIAL. |
| FT | TRANSMEM | 87 | 107 | POTENTIAL. |
| SQ | SEQUENCE | 307 AA; | 33921 MW; | E66734B84410996D CRC64; |

Query Match 24.9%, Score 51.5; DB 1; Length 307;
 Best Local Similarity 32.5%; Pred. No. 24;
 Matches 13; Conservative 8; Mismatches 14; Indels 5; Gaps 1;

OY 5 IELTGMVSVILGAVASHIDNYOPPEOSASVQ 39
 DB 87 IELTGMVSVILGAVASHIDNYOPPEOSASVQ 126

RESULT 13

TCR_STRAG STANDARD; PRT; 458 AA.
 ID TCR_STRAG STANDARD; PRT; 458 AA.
 AC P13924;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Tetracycline resistance protein.
 GN TET.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus
 OX NCBI_TaxID=1311;
 RN [1]
 RM SEQUENCE FROM N.A.
 RX MEDLINE-90016790; PubMed-2677995;
 RA van der Lelie D., Bron S., Venema G., Oskam L.;
 RT "Similarity of minus origins of replication and flanking open reading
 RL frames of plasmids PUB10, pTR913 and pMW158."
 CC Nucleic Acids Res. 17:7283-7294(1989).

CC -1 FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1 SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X15669; CAA33712.1; -

DR PIR: C25599; YTSOG.
 DR InterPro: IPR001411; TCR_FetB.

DR PRINTS: PR01036; TCR_FetB.

KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.

FT TRANSMEM 12 33 POTENTIAL.
 FT TRANSMEM 81 100 POTENTIAL.
 FT TRANSMEM 111 129 POTENTIAL.
 FT TRANSMEM 140 162 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 223 240 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 346 365 POTENTIAL.
 FT TRANSMEM 432 451 POTENTIAL.
 SQ SEQUENCE 458 AA; 50006 MW; ADZ01E7CA199993 CRC64;

Query Match 24.6%; Score 51; DB 1; Length 458;
 Best Local Similarity 28.9%; Pred. No. 41;
 Matches 11; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

OY 4 LIELTGMVSVILGAVASHIDNYOPPEOSASVQHK 41
 DB 343 LIELTGMVSVILGAVASHIDNYOPPEOSASVQHK 380

RESULT 14
 ID YHV2_LACHE STANDARD; PRT; 104 AA.
 AC P22295;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 11.8 kDa protein in HIV 5' region precursor (ORF2).
 OS Lactobacillus helveticus.
 CC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 OX NCBI_TaxID=1587;
 RN [1]
 RM SEQUENCE FROM N.A.
 RP SPRAIN-481;
 RX MEDLINE-91035244; PubMed-2228964;
 RA Joergers M.C., Kleenhammer T.R.;
 RT "Cloning, expression, and nucleotide sequence of the Lactobacillus
 RL J. Bacteriol. 172:6339-6347(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M59360; AAA63273.1; -
 DR PIR: B37145; B37145.
 KW Hypothetical protein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 104 HYPOTHETICAL 11.8 KDA PROTEIN.
 SQ SEQUENCE 104 AA; 11809 MW; 399818546204C9B CRC64;

Query Match 24.4%; Score 50.5; DB 1; Length 104;
 Best Local Similarity 47.8%; Pred. No. 11;
 Matches 11; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

OY 4 LIELTGMVSVILGAVASHIDNYOPPEOSASVQHK 41
 DB 9 LIALAFWISVSV---GLISHV 28

RESULT 15

CCME_HAEIN STANDARD; PRT; 173 AA.
 ID CCME_HAEIN STANDARD; PRT; 173 AA.
 AC PA5036;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome c-type biogenesis protein ccme.
 GN CCME OR H11093.
 OS Haemophilus influenzae.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RM SEQUENCE FROM N.A.
 RP STRAIN-Rd / RW20 / ATCC 51907;
 RX MEDLINE-95350630; PubMed-7542800;
 RA Fritschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bull C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glöck A., Kelley J.M.,
 RA Welterman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Ullrich J., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae

```

RT Rd.":
RL Science 269:496-512(1995).
CC -1- FUNCTION: REQUIRED FOR THE BIOGENESIS OF C-TYPE CYTOCHROMES.
CC POSSIBLE SUBUNIT OF A HEME LYASE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MOSTLY PERIPLASMIC; ANCHORED IN THE INNER
CC MEMBRANE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE CME/CYCJ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32789; AAC22750.1; -.
CC TIGR: H11093; -.
DR InterPro: IPR004329; CME.
DR Pfam: PF03100; CME; 1.
KW Cytochrome c-type biogenesis; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT DOMAIN 30 173 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 173 AA; 19149 MW; 0505AC6B99709156 CRC64;

Query Match 24.4%; Score 50.5; DB 1; Length 173;
Best Local Similarity 40.5%; Pred. No. 18;
Matches 15; Conservative 4; Mismatches 13; Indels 5; Gaps 2;

QY 7 LTGMIVLVSVIILGVASHDNQPPQASVQ--HK 41
II II : : I I : I I I I : I I
Db 116 LTQPTVLTATEVL--AKHDENVPPPELGKMKQVHK 149

```

Search completed: February 6, 2003, 21:38:32
 Job time : 13 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:36:40 : Search time 29 Seconds
(without alignments)
291.308 Million cell updates/sec

Title: US-09-828-456-2

Perfect score: 207

Sequence: 1 MNRLIELTGMVLVSVILL.....VASHIDNYQPEQSASVQHK 41

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_plant:*
- 10: sp_protent:*
- 11: sp_virus:*
- 12: sp_vertebrate:*
- 13: sp_virus:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|-----------------------|
| 1 | 200 | 96.6 | 41 | 08X361 | 08X361 escherichia |
| 2 | 64 | 30.9 | 296 | 17 08TYS7 | 08TYS7 methanopyru |
| 3 | 63.5 | 30.7 | 254 | 2 091BY3 | 091BY3 shewanella |
| 4 | 58 | 28.0 | 332 | 16 08RM15 | 08RM15 rhizobium 1 |
| 5 | 58 | 28.0 | 344 | 10 08RMH8 | 08RMH8 arabidopsis |
| 6 | 57 | 27.5 | 440 | 2 052721 | 052721 rhodospirillum |
| 7 | 57 | 27.5 | 460 | 2 09T906 | 09T906 rhodospirillum |
| 8 | 57 | 27.5 | 826 | 3 09UVB5 | 09UVB5 candida alb |
| 9 | 56.5 | 27.3 | 309 | 13 08UWE8 | 08UWE8 tetradon n |
| 10 | 56.5 | 27.3 | 318 | 13 08QFS3 | 08QFS3 tetradon n |
| 11 | 56 | 27.1 | 490 | 2 09XBH1 | 09XBH1 prausserella |
| 12 | 55.5 | 26.8 | 623 | 16 08XN66 | 08XN66 clostridium |
| 13 | 55 | 26.6 | 219 | 16 097P36 | 097P36 streptococ |
| 14 | 55 | 26.6 | 320 | 16 09CFR7 | 09CFR7 lactococcus |
| 15 | 54.5 | 26.3 | 183 | 2 09JFR2 | 09JFR2 rhodococcus |
| 16 | 54 | 26.1 | 374 | 17 097X02 | 097X02 sulfolobus |

| | | | | | |
|----|------|------|------|-----------|---------------------|
| 17 | 54 | 26.1 | 2083 | 11 060997 | 060997 mus musculu |
| 18 | 53.5 | 25.8 | 65 | 16 08ZFE6 | 08ZFE6 yersinia pe |
| 19 | 53.5 | 25.8 | 206 | 2 09S446 | 09S446 staphylococ |
| 20 | 53.5 | 25.8 | 206 | 16 099RA7 | 099RA7 staphylococ |
| 21 | 53.5 | 25.8 | 344 | 16 08YEH7 | 08YEH7 brucella me |
| 22 | 53.5 | 25.8 | 379 | 8 034846 | 034846 katharina t |
| 23 | 53.5 | 25.8 | 432 | 13 098983 | 098983 raja erinac |
| 24 | 53 | 25.6 | 101 | 8 09BAC2 | 09BAC2 euglena myx |
| 25 | 53 | 25.6 | 553 | 13 08UYX3 | 08UYX3 gallus gall |
| 26 | 52.5 | 25.4 | 102 | 17 09H1Z6 | 09H1Z6 thermoplasm |
| 27 | 52.5 | 25.4 | 579 | 16 08UDV9 | 08UDV9 agrobacteri |
| 28 | 52.5 | 25.4 | 704 | 5 09GY56 | 09GY56 leishmania |
| 29 | 52 | 25.1 | 179 | 16 09CPM8 | 09CPM8 pasteurella |
| 30 | 52 | 25.1 | 459 | 5 09VX42 | 09VX42 drosophila |
| 31 | 52 | 25.1 | 506 | 10 0948L0 | 0948L0 oryza sativ |
| 32 | 52 | 25.1 | 506 | 10 0944W2 | 0944W2 oryza sativ |
| 33 | 52 | 25.1 | 739 | 12 08V717 | 08V717 ft virus. o |
| 34 | 52 | 25.1 | 1033 | 16 08XOM3 | 08XOM3 ralsstonia s |
| 35 | 52 | 25.1 | 1125 | 12 09L1I6 | 09L1I6 turkey astr |
| 36 | 52 | 25.1 | 5038 | 11 090YX7 | 090YX7 mus musculu |
| 37 | 51.5 | 24.9 | 90 | 12 08V201 | 08V201 emilliana hu |
| 38 | 51.5 | 24.9 | 90 | 12 08V200 | 08V200 emilliana hu |
| 39 | 51.5 | 24.9 | 92 | 12 08V1Z9 | 08V1Z9 emilliana hu |
| 40 | 51.5 | 24.9 | 92 | 12 08V1Z7 | 08V1Z7 emilliana hu |
| 41 | 51.5 | 24.9 | 94 | 12 08V204 | 08V204 emilliana hu |
| 42 | 51.5 | 24.9 | 94 | 12 08V202 | 08V202 emilliana hu |
| 43 | 51.5 | 24.9 | 94 | 12 08V1Z8 | 08V1Z8 emilliana hu |
| 44 | 51.5 | 24.9 | 94 | 12 08V1Z6 | 08V1Z6 emilliana hu |
| 45 | 51.5 | 24.9 | 199 | 16 09ZNN6 | 09ZNN6 rhizobium m |

ALIGNMENTS

| | | | | | |
|--|-----------------|------------------|-----------|------------|--|
| RESULT 1 | | | | | |
| 08X361 | | PRELIMINARY: | PRT: | 41 AA. | |
| AC 08X361: | | | | | |
| DT 01-MAR-2002 (TREMBLrel. 20, Created) | | | | | |
| DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update) | | | | | |
| DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update) | | | | | |
| DE Beta-lactam resistance protein. | | | | | |
| GN ECS2333. | | | | | |
| OS Escherichia coli O157:H7. | | | | | |
| OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | | | | |
| OC Escherichia. | | | | | |
| OX NCBI_TaxID=83334; | | | | | |
| RN [1] | | | | | |
| RP SEQUENCE FROM N.A. | | | | | |
| RC STRAIN=O157:H7 / RIMD 0509952; | | | | | |
| RX MEDLINE=21156231; PubMed=11258796; | | | | | |
| RA Hayashi T., Makino K., Chitishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kihara S., Shiba T., Hattori M., Shinagawa H.; | | | | | |
| RA *Complete genome sequence of enterohemorrhagic Escherichia coli | | | | | |
| RT O157:H7 and genomic comparison with a laboratory strain K-12.*; | | | | | |
| RL DNA Res. 8:11-22(2001). | | | | | |
| DR EMBL, AP002558; BAB35756.1; - | | | | | |
| SQ SEQUENCE 41 AA; 4600 MW; D384EBB0167AA46B CRC64; | | | | | |
| Query Match | 96.6%; | Score 200; | DB 16; | Length 41; | |
| Best Local Similarity | 95.1%; | Pred. No. 7e-20; | | | |
| Matches 39; | Conservative 1; | Mismatches 1; | Indels 0; | Gaps 0; | |
| OY 1 MNRLIELTGMVLVSVILLGVASHIDNYQPEQSASVQHK 41 | | | | | |
| DB 1 MNRLIELTGMVLVSVILLGVASHIDNYQPEQSTVQHK 41 | | | | | |
| RESULT 2 | | | | | |
| 08TYS7 | | PRELIMINARY: | PRT: | 296 AA. | |
| ID 08TYS7 | | | | | |

| | | | |
|-----------------------|---|-----------------------|---|
| AC | 08RT57; | DT | 01-OCT-2001 (TReMBLrel. 18, last sequence update) |
| DT | 01-JUN-2002 (TReMBLrel. 21, Created) | DT | 01-MAR-2002 (TReMBLrel. 20, last annotation update) |
| DT | 01-JUN-2002 (TReMBLrel. 21, last sequence update) | DE | Hydroxymuconic semialdehyde hydrolase. |
| DT | 01-JUN-2002 (TReMBLrel. 21, last annotation update) | GN | MR2612. |
| DE | Predicted membrane protein. | OS | Rhizobium loti (Mesorhizobium loti). |
| GN | ME035. | OC | Bacteria; Proteobacteria; alpha-subdivision; Rhizobiaceae group. |
| OC | Methanopyrus kandleri. | OC | Phyllobacteriaceae; Mesorhizobium. |
| OC | Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae; | OC | NCBI_TaxID=381; |
| OC | Methanopyrus. | OX | NCBI_TaxID=381; |
| OX | NCBI_TaxID=2320; | RP | SEQUENCE FROM N.A. |
| RN | [1] | RC | STRAIN-MAF303099; |
| RP | SEQUENCE FROM N.A. | RX | MEDLINE-21082930; PubMed-11214968; |
| RC | STRAIN-AV19 / DSM 6324 / JCM 9639; | RA | kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., |
| RX | MEDLINE-21927647; PubMed-11930014; | RA | Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T., |
| RA | Slesarev A.I., Mezhevaia K.V., Makarova K.S., Polushin N.N., | RA | Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., |
| RA | Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L., | RA | Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M., |
| RA | Natale D.A., Rogozin I.B., Talusov R.L., Wolf Y.I., Stetter K.O., | RT | Takeuchi C., Yamada M., Tabata S.; |
| RT | "The complete genome of hyperthermophile Methanopyrus kandleri AV19 | RT | "Complete genome structure of the nitrogen-fixing symbiotic bacterium |
| RT | and monophyly of archaeal methanogens." | RT | Mesorhizobium loti." |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002). | RL | DNA Res. 7:331-338 (2000). |
| DR | EMBL; AEO10320; AMO1432.1; | DR | EMBL; AP003000; BAB49701.1; |
| KM | Complete proteome. | DR | InterPro: IPR000073; Abhydrolase. |
| SO | SEQUENCE 296 AA; 32077 MW; 5FC5C8A123866BA7 CRC64; | DR | InterPro: IPR003089; AB_hydrolase. |
| Query Match | 30.9%; Score 64; DB 17; Length 296; | DR | InterPro: IPR000639; Epox_hydrolase. |
| Best Local Similarity | 35.3%; Pred. No. 1.3; | DR | InterPro: IPR00379; Ser_estr_s-site. |
| Matches | 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0; | DR | Pfam: PF00561; abhydrolase; 1. |
| OY | 6 ELTGMIVLVSVILGVASHNDNQPEQASVQ 39 | DR | PRINTS: PR00412; EPOXYDRASE. |
| Db | 170 ELTGMIVLVSVGVAVTATIDEVPEPEYADLK 203 | KM | Hydrolase; Complete proteome. |
| RESULT 3 | | SO | SEQUENCE 332 AA; 35933 MW; B33885190AF59003 CRC64; |
| 09LBY3 | | Query Match | 28.0%; Score 58; DB 16; Length 332; |
| ID | 09LBY3 | Best Local Similarity | 33.3%; Pred. No. 9.9; |
| AC | 09LBY3; | Matches | 14; Conservative 11; Mismatches 11; Indels 6; Gaps 2; |
| DT | 01-OCT-2000 (TReMBLrel. 15, Created) | OY | 1 MNRLELTGW----IYLVSVILGV--ASHDNDQPEPQSA 36 |
| DT | 01-OCT-2000 (TReMBLrel. 15, last sequence update) | Db | 1 MRLRLIVGWLAVTLVVLVVLVLAIFRGAALRENSPEEA 42 |
| DT | 01-JUN-2002 (TReMBLrel. 21, last annotation update) | RESULT 5 | |
| DE | Ouinol oxidase subunit III. | 08RWH8 | |
| GN | CYOA. | ID | 08RWH8 |
| OS | Shewanella violacea. | AC | 08RWH8; |
| OC | Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae; | DT | 01-JUN-2002 (TReMBLrel. 21, Created) |
| OC | Shewanella. | DT | 01-JUN-2002 (TReMBLrel. 21, last sequence update) |
| OX | NCBI_TaxID=60217; | DT | 01-JUN-2002 (TReMBLrel. 21, last annotation update) |
| RN | [1] | DE | Hypothetical 37.4 kDa protein. |
| RP | SEQUENCE FROM N.A. | GN | AT5G11960. |
| RA | Qureshi M.H., Kato C., Nakasone K., Yamada M., Horikoshi K.; | OS | Arabidopsis thaliana (Mouse-ear cress). |
| RT | "Pressure-regulation of a membrane-bound quinol oxidase in a deep-sea | OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; |
| RT | plezophilic bacterium, Shewanella violacea." | OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; |
| RL | Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases. | OC | eurocots II; Brassicales; Brassicaceae; Arabidopsis. |
| DR | EMBL; AB033827; BAA94864.1; | OX | NCBI_TaxID=3702; |
| DR | HSSP: P18400; ICYW. | RN | [1] |
| DR | InterPro: IPR001505; Copper_Cua. | RP | SEQUENCE FROM N.A. |
| DR | Prodrom: PD000131; Copper_Cua; 1. | RA | Soutlwick A., Karlun-Neumann G., Nguyen M., Lam B., Miranda M., |
| SO | SEQUENCE 254 AA; 28657 MW; F1870460580ACED CRC64; | RA | Palm C.J., Bowser L., Jones T., Banh J., Carrinici P., Chen H., |
| Query Match | 30.7%; Score 63.5; DB 2; Length 254; | RA | Chen R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J., |
| Best Local Similarity | 35.0%; Pred. No. 1.3; | RA | Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., |
| Matches | 14; Conservative 9; Mismatches 12; Indels 5; Gaps 1; | RA | Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., |
| OY | 2 NRIELTGMIVLVSVILGV----ASHDNDQPEPQSA 36 | RA | Becker J., Theologis A., Davis R.W.; |
| Db | 35 SKTIEIVWVPIVIVIVIGVITWTSTQDLDPRKPLDHEA 74 | RL | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases. |
| RESULT 4 | | DR | EMBL; AY093077; AAM13076.1; |
| 098115 | | KM | Hypothetical protein. |
| AC | 098115 | SO | SEQUENCE 344 AA; 37388 MW; 0D941E9A8FC177EA CRC64; |
| ID | 098115 | Query Match | 28.0%; Score 58; DB 10; Length 344; |
| DT | 01-OCT-2001 (TReMBLrel. 18, Created) | Best Local Similarity | 34.5%; Pred. No. 10; |
| Matches | 10; Conservative 10; Mismatches 9; Indels 0; Gaps 0; | OY | 3 RLIELTGMIVLVSVILGVASHDNDQY 31 |
| Db | 170 ELTGMIVLVSVGVAVTATIDEVPEPEYADLK 203 | Db | 1 MRLRLIVGWLAVTLVVLVVLVLAIFRGAALRENSPEEA 42 |

```

Db 265 RLLILGLMLMLGVLLVTSRLRLHLP 293

RESULT 6
ID 052721 PRELIMINARY; PRT: 440 AA.
AC 052721;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Sensor kinase.
GN REGB.
OS Rhodospirillum rubrum (Rhodospirillum rubrum)
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;
OC Rhodospirillum.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB1003;
RX MEDLINE=95095926; PubMed=8002581;
RA Mosley C.S., Suzuki J.Y., Bauer C.E.;
RT "Identification and molecular genetic characterization of a sensor
RT kinase responsible for coordinately regulating light harvesting and
RT reaction center gene expression in response to anaerobiosis [published
RT erratum appears in J Bacteriol 1995 Jun;177(11):3359].";
RL J. Bacteriol. 176:7566-7573(1994).
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSOR TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: L35179; AAA80144.1; -.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR004359; His_kinA.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; signal; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; HisKA; 1.
DR KINASE: Phosphorylation; Sensory transduction; Transferase.
SO SEQUENCE 440 AA; 47410 MW; 8BE58AA65C318C77 CRC64;

Query Match 27.5%; Score 57; DB 2; Length 440;
Best Local Similarity 40.9%; Pred. No. 18;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 LIELTGWLTVLVSVYLLGVASH 25
Db 146 MIEFGSWAIVGVIFLGAVAH 167

RESULT 7
ID 09L906 PRELIMINARY; PRT: 460 AA.
AC 09L906;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sensor kinase.
GN REGB.
OS Rhodospirillum rubrum (Rhodospirillum rubrum)
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;
OC Rhodospirillum.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB1003;
RX MEDLINE=20100775; PubMed=10633119;
RA Chen W., Jager A., Klug G.;
RT "Correction of the DNA sequence of the regB gene of Rhodospirillum
RT rubrum with implications for the membrane topology of the sensor
RT kinase RegB.";
RL J. Bacteriol. 182:818-820(2000).
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSOR TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AF189160; AA01775.3; -.

Db 166 MIEFGSWAIVGVIFLGAVAH 187

Query Match 27.5%; Score 57; DB 2; Length 460;
Best Local Similarity 40.9%; Pred. No. 19;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 LIELTGWLTVLVSVYLLGVASH 25
Db 166 MIEFGSWAIVGVIFLGAVAH 187

RESULT 8
ID 09UVB5 PRELIMINARY; PRT: 826 AA.
AC 09UVB5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Protein mannosyltransferase.
GN PMT6.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1006;
RX MEDLINE=20270133; PubMed=10809683;
RA Timpel C., Zink S., Strahl-Bolsinger S., Schroppe K., Ernst J.;
RT "Morphogenesis, adhesive properties, and antifungal resistance depend
RT on the pmt6 protein mannosyltransferase in the fungal pathogen Candida
RT albicans.";
RL J. Bacteriol. 182:3063-3071(2000).
DR EMBL: AF104916; AAFL6867.1; -.
DR InterPro: IPR003608; MTR.
DR InterPro: IPR003424; PMT.
DR Pfam: PF02815; MTR; 3.
DR Pfam: PF02366; PMT; 1.
DR SMART: SM00472; MTR; 3.
DR GLYCOSYLTRANSFERASE: Transferase.
SO SEQUENCE 826 AA; 94294 MW; B3DE7538C1F8EB7 CRC64;

Query Match 27.5%; Score 57; DB 3; Length 826;
Best Local Similarity 35.7%; Pred. No. 36;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 4 LIELTGWLTVLVSVYLLGVASHINDYOP 31
Db 729 LLPLAGWVLAHVPLLMGRVKKYLLHLYVP 756

RESULT 9
ID 08UWE8 PRELIMINARY; PRT: 309 AA.
AC 08UWE8;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE AF290615-like protein.
GN AF290615-LIKE.
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```

| | |
|-----------|---|
| CC | Acanthomorphia: Acanthopteriygii; Percomorpha; Tetraodontiformes; |
| OC | Tetraodontidae: Tetraodon. |
| OX | NCBI_TaxID=99883; |
| OX | NCBI_TaxID=99883; |
| RP | SEQUENCE FROM N.A. |
| RT | Analysis of 148 kb of genomic DNA of Tetraodon nigroviridis covering |
| RT | an amyase gene family. " |
| RL | Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases. |
| RL | EMBL; AJ308233; CAC87122.1; " |
| SQ | SEQUENCE 309 AA; 35327 MW; AF93C122AF1FE094 CRC64; |
| QY | Query Match 27.3%; Score 56.5; DB 13; Length 309; |
| Db | Best Local Similarity 40.5%; Pred. No. 15; |
| Matches | 15; Conservative 7; Mismatches 6; Indels 9; Gaps 2; |
| RESULT 10 | |
| ID | Q8QFS3 PRELIMINARY; PRT; 318 AA. |
| AC | Q8QFS3; |
| DT | 01-JUN-2002 (TREMBLrel. 21, Created) |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last sequence update) |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) |
| DE | H01CJC protein. |
| GN | H01CJC. |
| OS | Tetraodon nigroviridis (Green puffer). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Actinopterygii; Neopteleostei; Teleostei; Euteleostei; Neoteleostei; |
| OC | Acanthomorphia; Acanthopteriygii; Percomorpha; Tetraodontiformes; |
| OC | Tetraodontidae; Tetraodon. |
| OX | NCBI_TaxID=99883; |
| OX | NCBI_TaxID=99883; |
| RP | SEQUENCE FROM N.A. |
| RT | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | EMBL; AJ427223; CAD20261.1; " |
| SQ | SEQUENCE 318 AA; 36409 MW; 4F2952DC77005A CRC64; |
| QY | Query Match 27.3%; Score 56.5; DB 13; Length 318; |
| Db | Best Local Similarity 40.5%; Pred. No. 15; |
| Matches | 15; Conservative 7; Mismatches 6; Indels 9; Gaps 2; |
| RESULT 11 | |
| ID | Q9XBMI PRELIMINARY; PRT; 490 AA. |
| AC | Q9XBMI; |
| DT | 01-NOV-1999 (TREMBLrel. 12, Created) |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) |
| DE | Alkane 1-monooxygenase (EC 1.14.15.3). |
| GN | ALRB. |
| OS | Prasereella rugosa. |
| OC | Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; |
| OC | Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae; Prasereella. |
| OX | NCBI_TaxID=43354; |
| OX | NCBI_TaxID=43354; |
| RP | SEQUENCE FROM N.A. |
| RT | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | EMBL; AJ427223; CAD20261.1; " |
| SQ | SEQUENCE 318 AA; 36409 MW; 4F2952DC77005A CRC64; |
| QY | Query Match 27.3%; Score 56.5; DB 13; Length 318; |
| Db | Best Local Similarity 40.5%; Pred. No. 15; |
| Matches | 15; Conservative 7; Mismatches 6; Indels 9; Gaps 2; |
| RESULT 12 | |
| ID | Q8XN66 PRELIMINARY; PRT; 623 AA. |
| AC | Q8XN66; |
| DT | 01-MAR-2002 (TREMBLrel. 20, Created) |
| DT | 01-MAR-2002 (TREMBLrel. 20, Last sequence update) |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) |
| DE | Probable ABC transporter. |
| GN | CPE0472. |
| OS | Clostridium perfringens. |
| OC | Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia; |
| OC | Clostridiales; Clostridiaceae; Clostridium. |
| OX | NCBI_TaxID=1502; |
| OX | NCBI_TaxID=1502; |
| RP | SEQUENCE FROM N.A. |
| RT | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | EMBL; AJ427223; CAD20261.1; " |
| SQ | SEQUENCE 318 AA; 36409 MW; 4F2952DC77005A CRC64; |
| QY | Query Match 26.8%; Score 55.5; DB 16; Length 623; |
| Db | Best Local Similarity 29.7%; Pred. No. 42; |
| Matches | 11; Conservative 11; Mismatches 10; Indels 5; Gaps 1; |
| RESULT 13 | |
| ID | Q97P36 PRELIMINARY; PRT; 623 AA. |
| AC | Q97P36; |
| DT | 01-NOV-1999 (TREMBLrel. 12, Created) |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) |
| DE | Alkane 1-monooxygenase (EC 1.14.15.3). |
| GN | ALRB. |
| OS | Prasereella rugosa. |
| OC | Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; |
| OC | Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae; Prasereella. |
| OX | NCBI_TaxID=43354; |
| OX | NCBI_TaxID=43354; |
| RP | SEQUENCE FROM N.A. |
| RT | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | EMBL; AJ427223; CAD20261.1; " |
| SQ | SEQUENCE 318 AA; 36409 MW; 4F2952DC77005A CRC64; |
| QY | Query Match 26.8%; Score 55.5; DB 16; Length 623; |
| Db | Best Local Similarity 29.7%; Pred. No. 42; |
| Matches | 11; Conservative 11; Mismatches 10; Indels 5; Gaps 1; |
| RESULT 14 | |
| ID | Q97P36 PRELIMINARY; PRT; 623 AA. |
| AC | Q97P36; |
| DT | 01-NOV-1999 (TREMBLrel. 12, Created) |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) |
| DE | Alkane 1-monooxygenase (EC 1.14.15.3). |
| GN | ALRB. |
| OS | Prasereella rugosa. |
| OC | Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; |
| OC | Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae; Prasereella. |
| OX | NCBI_TaxID=43354; |
| OX | NCBI_TaxID=43354; |
| RP | SEQUENCE FROM N.A. |
| RT | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. |
| RL | EMBL; AJ427223; CAD20261.1; " |
| SQ | SEQUENCE 318 AA; 36409 MW; 4F2952DC77005A CRC64; |
| QY | Query Match 26.8%; Score 55.5; DB 16; Length 623; |
| Db | Best Local Similarity 29.7%; Pred. No. 42; |
| Matches | 11; Conservative 11; Mismatches 10; Indels 5; Gaps 1; |
| RESULT 15 | |
| ID | Q97P36 PRELIMINARY; PRT; 623 AA. |
| AC | Q97P36; |
| DT | 01-NOV-1999 (TREMBLrel. 12, Created) |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) |
| DE | Alkane 1-monooxygenase (EC 1.14.15.3). |
| GN | ALRB. |
| OS | Prasereella rugosa. |
| OC | Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; |
| OC | Actinomycetales; Pseudonocardiaceae; Pseudon |

| | |
|-----------|---|
| RX | STRAIN-NRRL B-2295; |
| RX | MEDLINE=211285526; PubMed=11207749; |
| RA | Smith T.H.M., Koehlisberger M., Witholt B., Van Belten J.B.; |
| RT | "Molecular screening for alkane hydroxylase genes in Gram-negative and Gram-positive bacteria."; |
| ST | Environ. Microbiol. 1:307-317(1999) |
| DR | EMBL: AJ009587; CAB51024.2; -. |
| DR | HSSP: P00268; SRXN. |
| DR | InterPro: IPR001225; FA_desaturase. |
| DR | InterPro: IPR004039; Rubredox. |
| DR | InterPro: IPR001052; Rubredoxin. |
| DR | Pfam: PF00487; FA_desaturase; 1. |
| DR | Pfam: PF00301; rubredoxin; 1. |
| DR | PRINTS: PR00163; RUBREDOXIN. |
| DR | ProDom: PD001610; Rubredoxin; 1. |
| DR | PROSITE: PS00202; RUBREDOXIN; 1. |
| KW | Monooxygenase; Oxidoreductase. |
| SQ | SEQUENCE 490 AA; 54700 MW; 83FAEA25D05B5848 CRC64; |
| OY | Query Match 27.1%; Score 56; DB 2; Length 490; |
| D6 | Best Local Similarity 36.8%; Pred. No. 28; |
| | Matches 14; Conservative 6; Mismatches 10; Indels 8; Gaps 1; |
| OY | 4 LIETGN-----IYLVSVILGVASHIDNYQPE 33 : : : : : : : : : : |
| D6 | 50 LREATGGMVLWGIPVILVYPPLDLVGLDRSNPD 87 |
| RESULT 12 | |
| OXNK66 | PRELIMINARY; PRT; 623 AA. |
| ID | OXNK66 |
| AC | OXNK66; |
| DT | 01-MAR-2002 (TrEMBLrel. 20, Created) |
| DT | 01-MAR-2002 (TrEMBLrel. 20, last sequence update) |
| DT | 01-JUN-2002 (TrEMBLrel. 21, last annotation update) |
| DE | Probable ABC transporter. |
| CN | CPE0472. |
| OS | Clostridium perfringens. |
| OC | Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia; |
| OC | Clostridiales; Clostridiaceae; Clostridium. |
| OX | NCBI_TaxID=1502; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN=13 / Type A; |
| RX | pubmed=11792842; |
| RA | Shimizu T., Ohtani K., Hirakawa H., Yamashita A., |
| RA | Shiba T., Ogasawara N., Hattori M., Kohara S., Hayashi H.; |
| RT | "Complete genome sequence of Clostridium perfringens, an anaerobic flesh-eater."; |
| ST | Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002). |
| DR | EMBL: AP003186; BAB80178.1; -. |
| DR | InterPro: IPR003593; AAA_ATPase. |
| DR | InterPro: IPR001140; ABCTranprtTM. |
| DR | InterPro: IPR003439; ABC_transportr. |
| DR | Pfam: PF00664; ABC_membrane; 1. |
| DR | Pfam: PF00005; ABC_tran; 1. |
| DR | ProDom: PD000006; ABC_transportr; 1. |
| DR | SMART: SM00382; AAA_1. |
| DR | PROSITE: PS00211; ABC_TRANSPORTER; 1. |
| KW | Complete proteome. |
| SQ | SEQUENCE 623 AA; 69773 MW; AFDS99795F257FA9 CRC64; |
| OY | Query Match 26.8%; Score 55.5; DB 16; Length 623; |
| D6 | Best Local Similarity 29.7%; Pred. No. 42; |
| | Matches 11; Conservative 11; Mismatches 10; Indels 5; Gaps 1; |
| OY | 10 WIVLVSVILGVASH-----IDNYQPPEQSASVQHK 41 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| D6 | 64 WLVIASIMLNITISVPISLINTYTSTKMSQTLOKK 100 |
| RESULT 13 | |
| 997P36 | |


```
ID 097P36 PRELIMINARY; PRT; 219 AA.
AC 097P36;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Type IV prepilin peptidase, putative.
GN SPI808.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4;
RX MEDLINE-21357209; PubMed-11463916;
RA Pettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Debey R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzaple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Anguioni S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
DR EMBL: AE007473; AAK75881.1; -.
DR TIGR: SPI808; -.
DR InterPro: IPR000045; Peptidase_C20.
DR Pfam: PF01478; Peptidase_C20; 1.
KW Complete proteome.
SQ SEQUENCE 219 AA; 24843 MW; 9FB487E094C90DF CRC64;

Query Match 26.6%; Score 55; DB 16; Length 219;
Best Local Similarity 40.9%; Pred. No. 16;
Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 4 LIELTGMIVLVSVILGVASH 25
Db 127 LIASGMVLVMSFLILGLIAH 148

RESULT 14
Q9CFA7 PRELIMINARY; PRT; 320 AA.
AC 09CFA7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein yqbf.
GN Y0BF OR LI1574.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-111403;
RX MEDLINE-21235186; PubMed-11337471;
RA Holcetin A., Wincker P., Manger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis 111403."
RN [1]
RP SEQUENCE FROM N.A.
RC EMBL: AE006388; AAK05672.1; -.
DR InterPro: IPR003089; AB_hydrolase.
DR InterPro: IPR000734; Lipase.
DR InterPro: IPR000379; Ser_estrs_site.
DR PRINTS: PRO0111; ABHYDROLASE.
DR PROSITE: PS00120; LIPASE_SER; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 320 AA; 35901 MW; 2E706FBDB803461A CRC64;
```

```
Query Match 26.6%; Score 55; DB 16; Length 320;
Best Local Similarity 37.9%; Pred. No. 24;
Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 1 MNRLIELTGMIVLVSVILGVASHIDNY 29
Db 1 MKRRKRMIVLVSVILGVGYFAIGNY 29

RESULT 15
Q93JR2 PRELIMINARY; PRT; 183 AA.
AC 093JR2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Alkane-1-monooxygenase (fragment).
GN ALKB.
OS Rhodococcus fascians.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-154-S;
RA Van Belen J.B., Smits T.H.M., Balada S.B., Witthol B.;
RT "Alkane hydroxylases in Gram-positive strains."
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ301873; CAC40958.1; -.
DR InterPro: IPR001225; FA_desaturase.
DR Pfam: PF00487; FA_desaturase; 1.
DR PRINTS: PRO1565; NEUROMEDINR.
KW Monooxygenase.
FT NON_TER 1
FT NON_TER 183
SQ SEQUENCE 183 AA; 21185 MW; 757649087E78C6D0 CRC64;

Query Match 26.3%; Score 54.5; DB 2; Length 183;
Best Local Similarity 25.5%; Pred. No. 15;
Matches 13; Conservative 13; Mismatches 12; Indels 13; Gaps 2;

QY 4 LIELTGMIVLVSVILGVASHIDNY-----QPEQASVQHK 41
Db 106 LIAVFGMIVLVSVILGVASHIDNY-----QPEQASVQHK 156
```

Search completed: February 6, 2003, 21:39:09
Job time : 33 secs